

Rapid Estimation of Earthquake Source and Ground-Motion Parameters for Earthquake Early Warning Using Data from a Single Three-Component Broadband or Strong-Motion Sensor

by M. Böse, T. Heaton, and E. Hauksson

Abstract We propose a new algorithm to rapidly determine earthquake source and ground-motion parameters for earthquake early warning (EEW). This algorithm uses the acceleration, velocity, and displacement waveforms of a single three-component broadband (BB) or strong-motion (SM) sensor to perform real-time earthquake/noise discrimination and near/far source classification. When an earthquake is detected, the algorithm estimates the moment magnitude M , epicentral distance Δ , and peak ground velocity (PGV) at the site of observation. The algorithm was constructed by using an artificial neural network (ANN) approach. Our training and test datasets consist of 2431 three-component SM and BB records of 161 crustal earthquakes in California, Japan, and Taiwan with $3.1 \leq M \leq 7.6$ at $\Delta \leq 115$ km. First estimates become available at $t_0 = 0.25$ s after the P pick and are regularly updated. We find that displacement and velocity waveforms are most relevant for the estimation of M and PGV, while acceleration is important for earthquake/noise discrimination. Including site corrections reduces the errors up to 10%. The estimates improve by an additional 10% if we use both the vertical and horizontal components of recorded ground motions. The uncertainties of the predicted parameters decrease with increasing time window length t_0 ; larger magnitude events show a slower decay of these uncertainties than small earthquakes. We compare our approach with the τ_c algorithm and find that our prediction errors are around 60% smaller. However, in general there is a limitation to the prediction accuracy an EEW system can provide if based on single-sensor observations.

Introduction

Earthquake early warning (EEW) techniques have improved significantly over the last decade, including both technological advances in real-time seismology and the development of algorithms for the rapid detection of possibly damaging earthquakes a few seconds to some tens of seconds before strong shaking occurs (Allen *et al.*, 2009). These algorithms require the seismic waveforms either from a single seismic sensor (the so-called on-site warning systems; e.g., Wu and Kanamori, 2005; Kanamori, 2005; Zollo *et al.*, 2006; Böse *et al.*, 2007) or from a seismic network or subnetwork (the so-called regional warning systems; e.g., Wu and Teng, 2002; Allen and Kanamori, 2003; Cua and Heaton, 2007). On-site and regional warning approaches deliver estimates of source and ground-motion parameters with different speed and accuracy (e.g., Kanamori, 2005).

Recently, Böse (2006) and Böse *et al.* (2008) developed an algorithm for EEW called PreSEIS that is based on artificial neural networks (ANNs). Artificial neural networks have several important features that make them attractive

tools for EEW. They allow for nonlinear mapping between the seismic waveforms recorded at one or more seismic sensors and the predicted source and ground-motion parameters at a user site. They do not require explicit formulations of relations, because they are completely data-driven and learn from examples or experience (similar to the human brain). They exhibit a high tolerance against noisy data, which is a common problem in real-time seismology, and they are computationally efficient, that is, very fast, which makes them applicable to real-time procedures such as EEW (Böse *et al.*, 2008).

PreSEIS fills the gap between on-site and regional warning methods. To estimate the source and ground-motion parameters of an ongoing earthquake, PreSEIS uses the seismic waveforms from multiple sensors in a seismic network without requiring that the seismic P wave has reached all of them yet. Nontriggered sensors provide important information about the source location by limiting the space of possible solutions (Horiuchi *et al.*, 2005; Böse *et al.*, 2008). The

continuous update of estimates allows the application of PreSEIS to large earthquakes with complex rupture evolution in which the largest slip along the fault does not necessarily occur close to the hypocenter, which is the point of rupture nucleation. PreSEIS is thus largely unaffected by whether or not the evolution of earthquake ruptures is predetermined at the beginning of the rupture process (e.g., [Olson and Allen, 2005](#); [Rydelek and Horiuchi, 2006](#); [Rydelek et al., 2007](#); [Yamada and Ide, 2008](#)).

PreSEIS was tested in several seismic-active regions around the world, including Istanbul ([Böse et al., 2008](#)), southern California ([Köhler et al., 2009](#)), Japan ([Köhler, 2010](#)), and Germany ([Hilbring et al., 2010](#)). The datasets used included (1) stochastic simulated strong-motion (SM) records, (2) sets of purely observed SM records, and (3) joint datasets of observed and simulated SM and broadband (BB) records. The results of all these studies demonstrated that ANNs are well suited for EEW ([Leach and Dowla, 1996](#)). Furthermore, they showed that the uncertainties in the predicted source and ground-motion parameters decrease with increasing length of the time window used, which causes a trade-off between the reliability of warnings and the remaining warning time until strong shaking occurs ([Böse et al., 2008](#)). These studies also revealed two major shortcomings of the PreSEIS algorithm. First, ANNs require large datasets for the training. Second, the PreSEIS algorithm is network-dependent, that is, once the ANNs have been trained for a particular seismic network or subnetwork, single sensors cannot be easily added or removed. While there are various methods to handle the problem of not-reporting stations at the time of an earthquake, for example, by the interpolation of values at neighboring stations, the robustness of this approach decreases if applied to less dense seismic networks.

To overcome these limitations, we propose the new PreSEIS On-site algorithm, which is solely based on single-sensor observations. For the training and testing of the algorithm, we use datasets of observed waveforms including BB and SM records from different tectonic regions (California, Japan, and Taiwan). We use these performance measures to make general conclusions about the expected uncertainties of estimates by an EEW algorithm, which uses the information from a limited time window of the P wave. We also compare our approach with the τ_c algorithm for EEW ([Kanamori, 2005](#)) and show that the prediction errors of PreSEIS On-site are around 60% smaller. However, in general there is a limitation to the prediction accuracy an EEW system can provide if based on single-sensor observations.

Method

Based on the observed waveforms at a single BB or SM sensor, PreSEIS On-site provides a rapid earthquake/noise discrimination, a near/far source classification, and estimates the moment magnitude M , the epicentral distance Δ , and the peak ground velocity (PGV) at the site of observation. All

estimates are updated with progressing time t_0 as more information about the earthquake becomes available. The principal approach is illustrated in Figure 1.

PreSEIS On-site uses the seismic acceleration, velocity, and displacement waveform time series, $\ddot{u}(t)$, $\dot{u}(t)$, and $u(t)$, recorded at a single three-component SM or BB sensor, obtained from the integration and differentiation of the recorded time series, respectively. The time series are parameterized by integrating the absolute amplitudes on component i ($i = \{EW, NS, UD\}$) over the time interval between the pick of the seismic P wave and a given time t_0 :

$$IAA_i \equiv \log_{10}(1 + \int_0^{t_0} |\ddot{u}_i(t)| dt), \quad (1)$$

$$IAV_i \equiv \log_{10}(1 + \int_0^{t_0} |\dot{u}_i(t)| dt), \quad (2)$$

and

$$IAD_i \equiv \log_{10}(1 + \int_0^{t_0} |u_i(t)| dt). \quad (3)$$

Taking the logarithmic values in equations (1) to (3) is important, because seismic amplitudes span several orders of magnitude and follow logarithmic distributions (e.g., [Yamada et al., 2009](#)). We add 1 to obtain positive values of IAA_i , IAV_i , and IAD_i . Equations (1) to (3) describe the envelope of the underlying waveform time series in a simplistic way.

Because the local soil conditions at the recording sites can lead to significant changes of the seismic wave amplitudes, and thus IAA_i , IAV_i , and IAD_i , it is important to consider these effects in our algorithm. We use the National Earthquake Hazard Reduction Program ([NEHRP; 1994](#)) V_{S30} site classification, that is, the average shear-wave velocity taken over the top 30 meters, as a simple proxy of local site conditions (Fig. 1). The V_{S30} value at a given station is a further input parameter to the ANNs. The ANNs have to find out themselves the relationship between this value and the changes of magnitudes, etcetera.

The discrimination between earthquakes and ambient noise (nonseismic events), as well as the near/far source classification, are essential parts of PreSEIS On-site. The latter is particularly needed for the real-time estimation of fault rupture length during large earthquakes ($M > 6.5$; e.g., [Yamada et al., 2007](#)). We define the following classification: an output of -1 is assigned if the pick was produced by noise; an output of $+1$ is assigned if the detected event is an earthquake with epicentral distance of Δ [km] $\leq 15M$; an output of 0 is assigned if the detected event is an earthquake with Δ [km] $> 15M$. The definition of this magnitude-dependent distance range is fairly arbitrary. It is driven by the observation that large-magnitude events can cause damaging ground shaking over larger areas than small earthquakes. While only the discrete values -1 , 0 , and $+1$ are assigned to the data for ANN learning, the output from the ANNs can be any rational

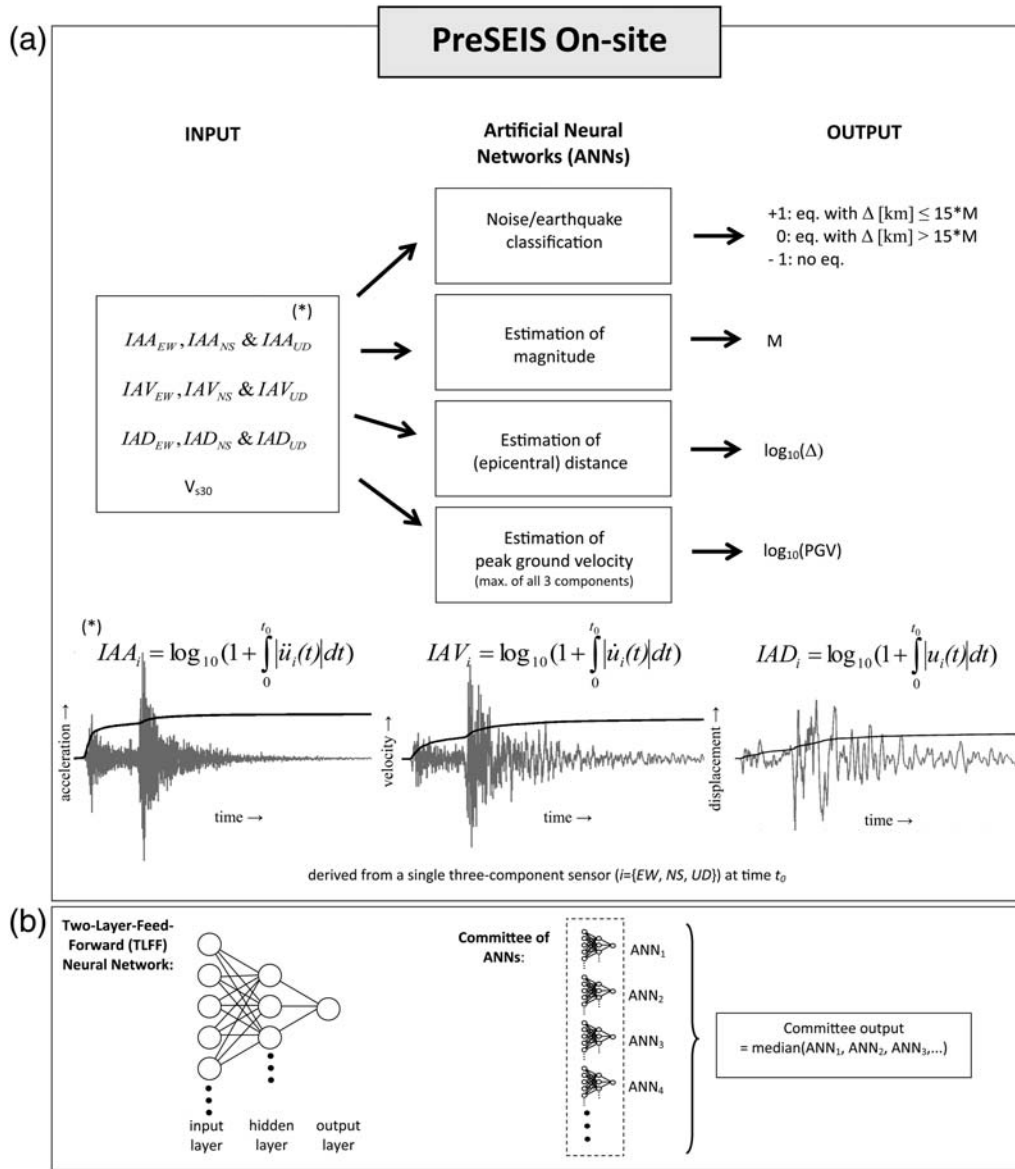


Figure 1. Principal approach of PreSEIS On-site. (a) The algorithm uses the logarithmic values of the integrated absolute amplitudes of acceleration, velocity, and displacement waveform time series, $\ddot{u}(t)$, $\dot{u}(t)$, and $u(t)$, at a single sensor, as well as V_{S30} site characterization. Outputs are (1) a simple earthquake/noise discrimination and near/far source classification, and estimates of (2) the moment magnitude M , (3) the epicentral distance Δ , and (4) the PGV. All estimates are updated with progressing time t_0 . (b) PreSEIS On-site uses two-layer-feed-forward (TLFF) neural networks composed of simple processing units arranged in input layers, hidden layers, and output layers that are connected to each other by a network of weighted links. Ten TLFF networks, which form a so-called committee, are trained on the same task (e.g., the prediction of M) using slightly different training datasets and weight initializations at the beginning of the training procedure; the median value taken over the outputs of all 10 TLFF networks defines the output of PreSEIS On-site.

number. For instance, 0.5 indicates that the event is probably an earthquake, and it is likely that $\Delta \approx 15M$.

As was the case for the former PreSEIS algorithm (Böse, 2006; Böse *et al.*, 2008), PreSEIS On-site uses two-layer-feed-forward (TLFF) neural networks for the mapping between the ground-motion observations and output parameters (Fig. 1, Appendix A). We train the ANNs for different time window lengths t_0 , ranging in intervals of 0.25 s from $t_0 = 0.25$ to $t_0 = 10.0$ s after the P pick. Each timestep has its own TLFF network (Böse *et al.*, 2008). Because we cannot

rule out that the training of the ANNs was insufficient (the optimization algorithm got stuck in a local minimum and the solution is not optimum), we construct so-called committees of ANNs (Fig. 1b; Bishop, 1995). All TLFF networks of one committee are trained with slightly different datasets, and the training starts with slightly different (randomly determined) weight initiations. In this paper we use committees of 10 TLFFs. The median over the outputs of all 10 TLFF networks defines the output of the PreSEIS On-site algorithm at a given time t_0 (Fig. 1b).

Data and Preprocessing

In this study we use a joint dataset of three-component BB and SM waveforms recorded by (1) the California Integrated Seismic Network (CISN), (2) the Japanese K-NET, and (3) the Taiwanese Strong-Motion Network (TSMIP). The datasets include the free-field records of small to large crustal earthquakes with different focal mechanisms and at different distances and soil conditions, as well as waveform time series of ambient noise.

Because EEW is most important for earthquakes causing significant levels of ground shaking, which primarily depends on magnitude M and distance Δ (aside from the local site conditions and the details of source radiation and wave propagation), we consider only records with Δ [km] $\leq 15M$. With this threshold, our database consists of 2431 three-component records of around 161 crustal earthquakes with $3.1 \leq M \leq 7.6$ and epicentral distances of up to 115 km (Fig. 2; Appendix B). In addition to the earthquake event data, we download continuous time series of ambient noise recorded at several CISN BB and SM stations. These data are used to train another ANN for the automatic discrimination of earthquakes and noise. If desired, the noise discrimination could be specific for a given station site, but we want to treat the problem more generally here.

For each record, we remove the trend and baseline and apply a gain correction. To obtain $\ddot{u}(t)$, $\dot{u}(t)$, and $u(t)$, we integrate and differentiate the data as appropriate. We apply a third-order causal Butterworth highpass filter with a corner

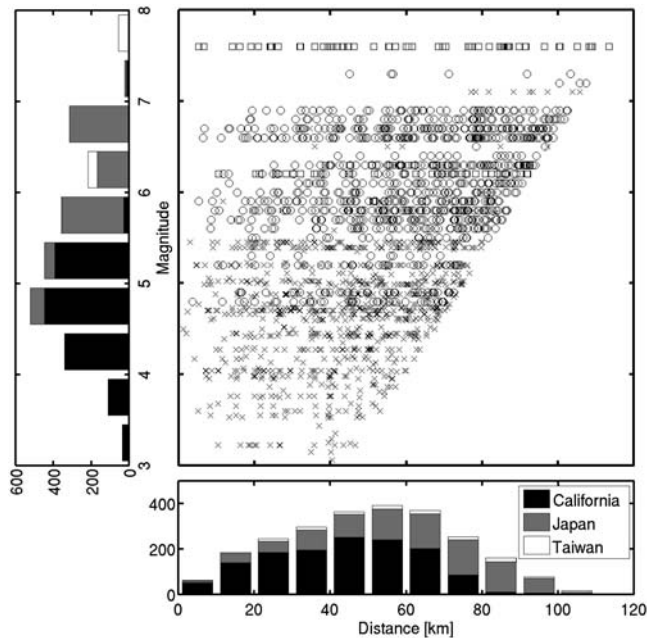


Figure 2. Histograms and distributions of magnitudes M and epicentral distances Δ of 2431 three-component BB and SM records of 161 crustal earthquakes from California, Japan, and Taiwan as used in this study. Because early warning is most important for earthquakes causing significant levels of ground shaking, we consider earthquakes with Δ [km] $\leq 15M$ only.

frequency at 0.075 Hz to remove long-period artifacts due to integration. The P -wave onset is automatically picked from the vertical velocity record (Allen, 1978). Then we use equations (1) to (3) to determine IAA_i , IAV_i , and IAD_i . The time windows for integration start from the P -pick and range in intervals of 0.25 s from $t_0 = 0.25$ to $t_0 = 10.0$ s.

Results

It is important for EEW to parameterize the recorded waveforms in a way that enables robust estimation of earthquake source and ground-motion parameters at a given time after rupture nucleation. Before analyzing the performance of PreSEIS On-site in more detail, we determine the time series (acceleration, velocity, or displacement) and components (horizontal or vertical) of the seismic data, or the combinations of these, that ensure the best mapping.

Displacement, Velocity, and/or Acceleration?

We train PreSEIS On-site with 90% of the available data (i.e., 2188 three-component records) adopting the earlier discussed training procedure for the ANNs (Appendix A). The training dataset is randomly selected, and we train a committee of 10 ANNs for the prediction of each output parameter. We repeat this procedure seven times, each time with different input information derived from the waveforms in the training dataset. We use the logarithmic values of the integrated absolute amplitudes of the acceleration, velocity, and displacement time series, IAA_i , IAV_i , and IAD_i (see equations (1) to (3)), and consider the following seven cases: usage of (1) IAD_i with V_{S30} ; (2) IAV_i with V_{S30} ; (3) IAA_i with V_{S30} ; (4) IAV_{UD} and IAD_{UD} with V_{S30} ; (5) IAV_i and IAD_i with V_{S30} ; (6) IAA_i , IAV_i , and IAD_i without V_{S30} ; and (7) IAA_i , IAV_i , and IAD_i with V_{S30} . Aside from case (4), we use all three components of ground motions ($i = \{EW, NS, UD\}$).

Independent from the input data chosen, the uncertainties of the predicted parameters (defined by the standard deviation σ of the Gaussian error distribution of observed and predicted output values for all 2431 records) decrease with increasing time t_0 ; that is, the longer we wait, the more reliable the estimates (Fig. 3). Later we show that the errors also depend on magnitude. The largest error reduction is observed within the first 2 to 3 s following the P -wave detection. Usually, the errors are largest if only IAA_i or IAV_i or IAD_i is used, while the best results are obtained when using a combination of the three. Including site corrections reduces the errors up to 10%. The estimates improve by an additional $\sim 10\%$ if we use both the vertical and horizontal components. The importance of each time series differs from output to output parameter (Fig. 3). While $u(t)$ and $\dot{u}(t)$ are most relevant for the estimation of M and PGV, $\ddot{u}(t)$ is more important for the earthquake/noise discrimination.

The percentage of misclassified events (false triggers) drops from 7% after $t_0 = 0.25$ s to 2.5% after $t_0 = 2$ s if

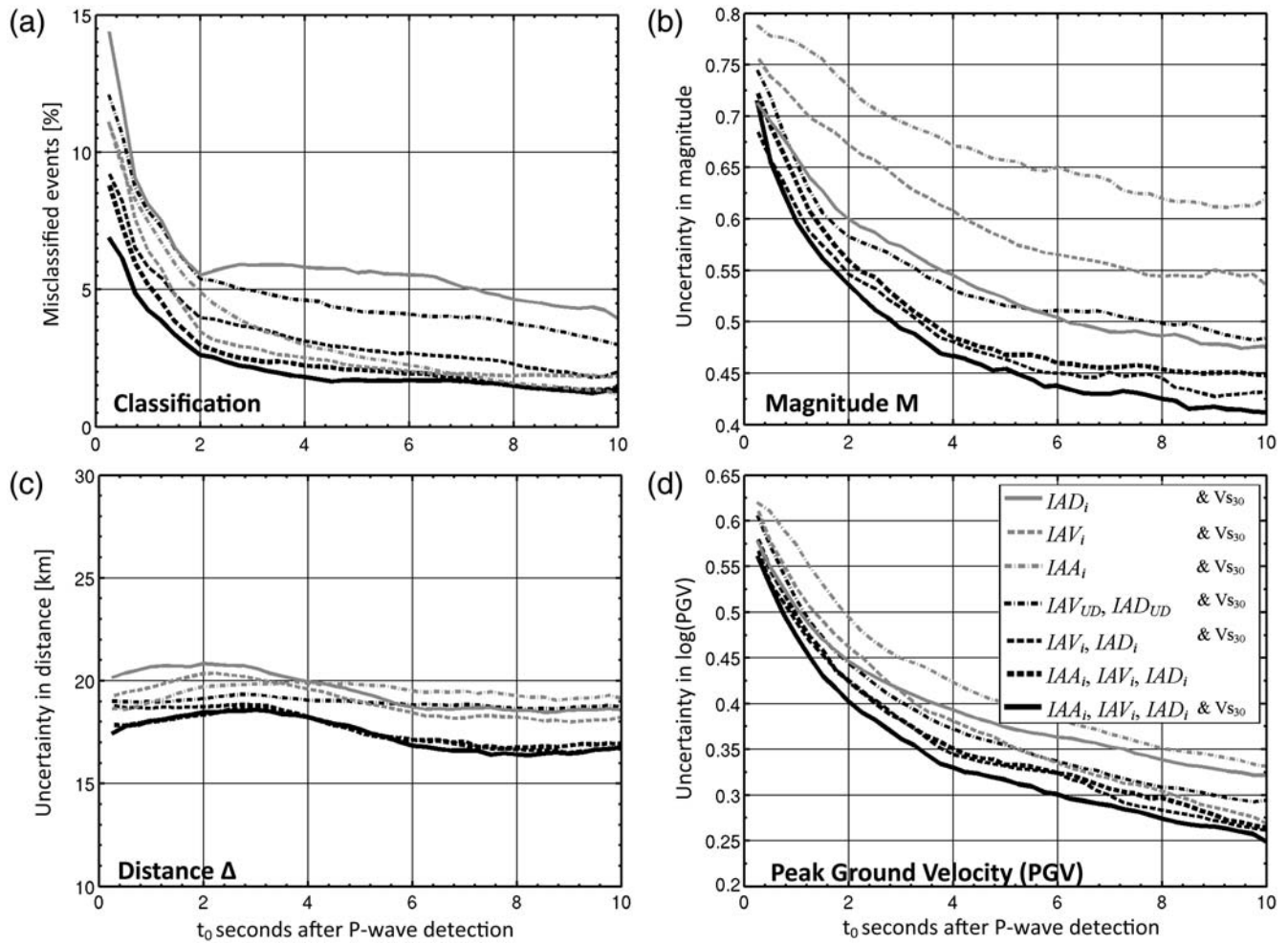


Figure 3. Smoothed errors in earthquake/noise classification and in estimated magnitude M , distance Δ , and peak ground velocity (PGV) as a function of time t_0 after the P -wave detection for different types of input information used. The best results are obtained when using three-component acceleration, velocity, and displacement data, IAA_i , IAV_i , and IAD_i , $i = \{EW, NS, UD\}$ (equations 1 to 3) with V_{S30} site corrections. High-frequency acceleration data are most important for classification; mid- and low-frequency velocity and displacement data play a dominant role for predicting the other parameters. Site corrections help improving the estimates by up to 10%, the usage of information derived from the horizontal components by an additional $\sim 10\%$. We show in Figure 4 that the errors also depend on M .

we use three-component acceleration, velocity, and displacement data with V_{S30} values (Fig. 3a). A larger time window does lead to a further small reduction of misclassified events ($\sim 1.5\%$).

For magnitude we determine a decrease in the uncertainties from around 0.7 units after $t_0 = 0.25$ s to 0.53 units after $t_0 = 2$ s and 0.5 units after $t_0 = 3$ s if we use IAA_i , IAV_i , and IAD_i with V_{S30} (Fig. 3b). A larger time window allows for a further reduction of these uncertainties. The uncertainties are only slightly smaller than if acceleration IAA_i is excluded.

There is almost no change in the uncertainty of epicentral distance Δ estimates with increasing t_0 . For example, if we use IAA_i , IAV_i , and IAD_i with V_{S30} , the uncertainty of Δ varies only slightly with values between 16 and 19 km (Fig. 3c). The distance estimates during the first $t_0 = 3$ s seem to be mainly based on the information derived from the acceleration waveform time series. After $t_0 = 3$ s, velocity and displacement start gaining in importance.

A rapid decrease in the uncertainties with increasing t_0 is also observed for PGV. The uncertainties are smallest for the combination of IAA_i , IAV_i , and IAD_i with V_{S30} (Fig. 3d). The errors in $\log(\text{PGV})$ decrease from ~ 0.55 after $t_0 = 0.25$ s to 0.4 cm/s after $t_0 = 2$ s and 0.35 cm/s after $t_0 = 3$ s. Acceleration IAA_i is the least important time series for this prediction, while displacement provides the most essential information about PGV up to $t_0 = 3$ s and is replaced by velocity for larger time windows.

Magnitude and Time Dependency of Prediction Errors

We have seen that PreSEIS On-site performs best if we combine three-component acceleration, velocity, and displacement data, IAA_i , IAV_i , and IAD_i , with V_{S30} site factors (Fig. 3). In the following we analyze the prediction

errors for this case in more detail. We focus on analyzing the magnitude errors $M_{\text{err}} = M_{\text{obs}} - M_{\text{pred}}$.

There are apparent trends of over- and underestimation of earthquake magnitudes (Fig. 4a). Magnitudes for small earthquakes with $M < 5.0$ tend to be overestimated, while magnitudes for large earthquakes are underestimated. The errors decrease exponentially with increasing time window length t_0 ; the larger the earthquake, the slower is the decay; that is, large events require more time to be recognized than small events.

To rule out that the magnitude and time dependencies of prediction errors in Figure 4a are biased by the training data or the training method of the ANNs, we repeat the training and testing of PreSEIS On-site several times using (1) different training datasets, (2) distinct error measures for optimization (mean absolute and sum squared errors), (3) different numbers of hidden neurons (Fig. 1b), and (4) consistent application of the same distance range of $\Delta \leq 100$ km to all events, rather than adopting a magnitude-dependent threshold of $\Delta [\text{km}] \leq 15M$.

The only factor that strongly affects the magnitude–time dependency of errors is the magnitude–frequency distribution of events in the dataset. While our original dataset contains events with almost uniform magnitude distribution (Fig. 2), earthquakes are commonly observed to follow power-law distributions, such as described by the Gutenberg–Richter relation (Gutenberg and Richter, 1944)

$$\log[M(N)] \propto -\beta M. \quad (4)$$

We use equation (4) to define a simple weighting function for the events in our original dataset and increase the

number of events of a given magnitude in the set accordingly before restarting the training procedure of the ANNs. To save computational time, we restrict this approach to events with $M \geq 4.0$.

The smallest errors for the uniformly distributed dataset are observed for events with $\sim M 5.0$ (Fig. 4a), and for the Gutenberg–Richter distributed set for events with $\sim M 4.0$ (Fig. 4b). This observation is not surprising because $\sim M 5.0$ and $M 4.0$ are the mean magnitudes of the two sets; we expect the best performance at these magnitudes after the optimization of the ANNs. Further, the magnitude–time dependencies of the prediction errors in Figure 4a,b are basically the same. Thus, the magnitude–time behavior of the prediction errors is not significantly affected by the composition of the dataset.

Broadband (BB) versus Strong-Motion (SM) Data

In general, EEW systems are based on BB or SM instrumentation. In this study, we use a dataset comprising the records of both sensor types and find clear differences in the performance of PreSEIS On-Site (Fig. 5). The standard deviations σ of the (Gaussian) error distributions in magnitude predictions are generally smaller for the BB than for the SM data, that is, the errors are smaller (Fig. 5a,b). The longer the time window t_0 of the P -waveform data used, the more accurate are the predictions. The smallest errors for BB are observed for earthquakes with $M \leq 5.5$, while for SM the errors are smallest for $4.5 \leq M \leq 6.5$. Note, however, that the BB and SM data in our set cover different magnitude ranges. Most BB records are from earthquakes with $3.1 \leq M \leq 6.0$, while most SM records are from events with

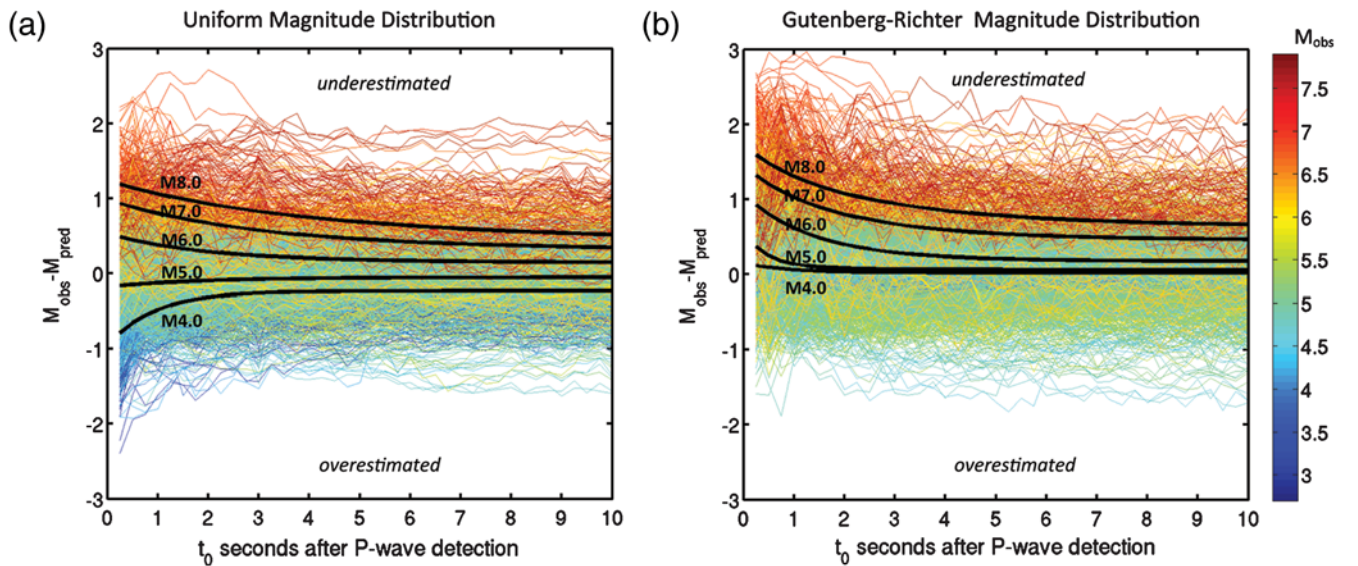


Figure 4. Magnitude prediction errors as a function of time t_0 and magnitude M for (a) a uniformly distributed dataset, and (b) a dataset distributed according to the Gutenberg–Richter power-law statistics; black lines show the mean values for $M 4.0$ to $M 8.0$. The results for $M 8.0$ are obtained from extrapolation. The main difference between (a) and (b) is a shift of the prediction errors to smaller and larger values, respectively; the magnitude–time dependence of the errors, however, remains almost the same. A small earthquake can be faster recognized than a large event.

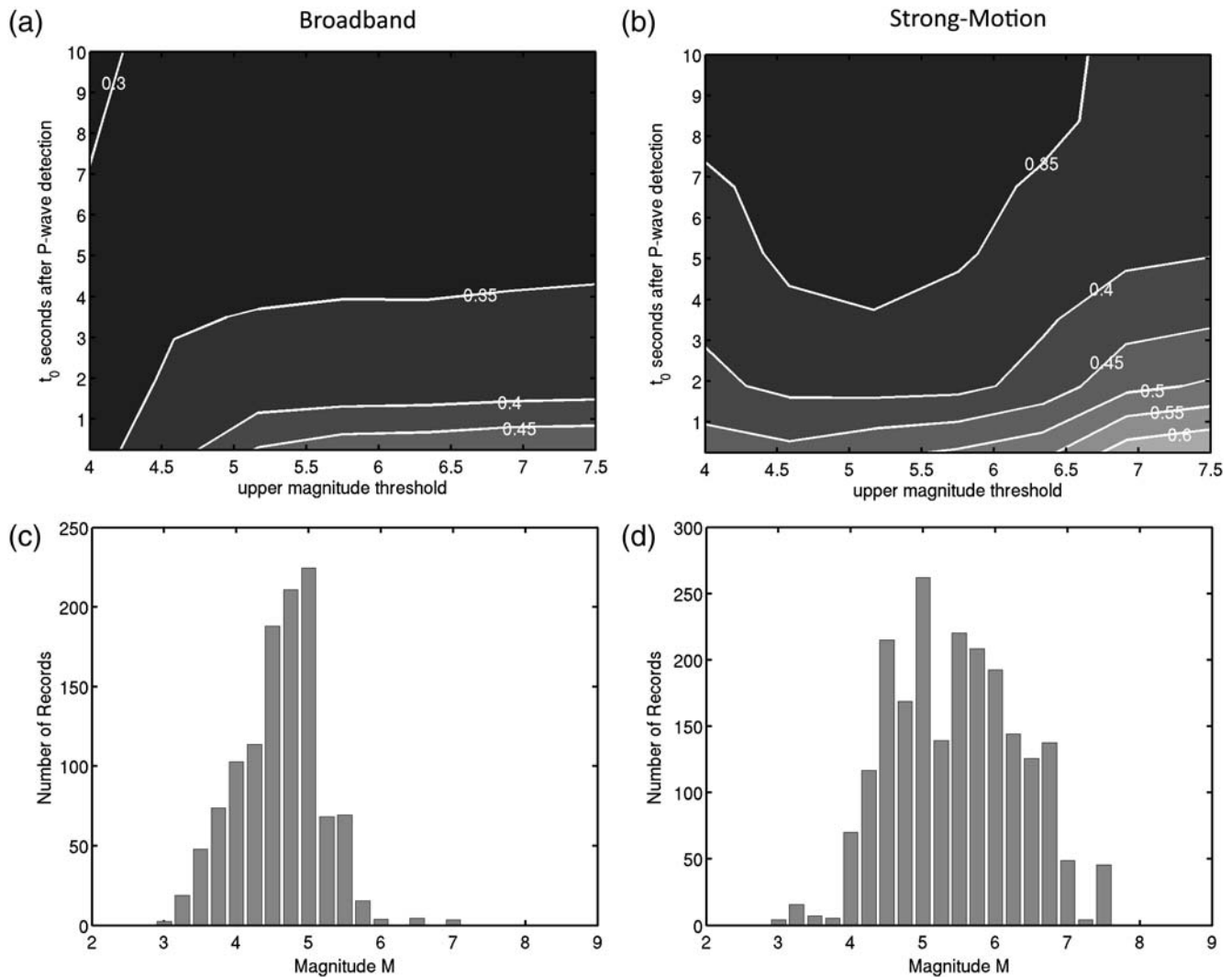


Figure 5. Comparison of magnitude prediction errors (standard deviation σ of error distributions) for (a) broadband (BB) and (b) strong-motion (SM) records as a function of magnitude M and time t_0 (top). The errors for the BB data are generally smaller than for SM; note, however, that most BB records are (c) $3.1 \leq M \leq 6.0$, while (d) $4.0 \leq M \leq 7.0$ for the SM data. The longer t_0 , the smaller the errors. The smallest errors for BB are observed for the smallest earthquakes ($M \leq 5.5$), while for SM the errors are smallest for $4.5 \leq M \leq 6.5$.

$4.0 \leq M \leq 7.0$ (Fig. 5c,d). Further, the dataset contains almost no records with clipped waveform amplitudes typically observed for BB sensors during large and close earthquakes. We therefore suspect that the better performance of the BB sensors for large-magnitude earthquakes in Figure 5 is some artifact caused by the limitation of the BB dataset.

Figure 5 suggests that BB sensors tend to perform better (due to the higher signal-to-noise ratio) during small- to moderate-sized earthquakes ($M < 6$) that occur more frequently, but usually do not cause significant damage. In contrast, because of their low gain, SM sensors outperform the BB sensors during very strong shaking.

Examples

In Figure 6 we analyze the performance of PreSEIS On-site, including the classification and the estimation of

M , PGV, and Δ within the first $t_0 = 10$ s after the P -wave detection for three example earthquakes from the test dataset; that is, these data were not used for the training of the ANNs: the 2010 M 4.1 Redlands (N34.00°/W117.18°/Z = 9 km) and 2010 M 5.4 Collins Valley earthquakes (N33.42°/W116.45°/Z = 14 km) in southern California, and the 2008 M 6.9 Miyagi earthquake (N39.03°/E140.88°/Z = 8 km) in Japan. For each of the three earthquakes, we randomly pick two stations at distinct epicentral distances Δ . Note that station IWT010 is very close to the city of Ōshū, Iwate Prefecture, which experienced damaging shaking during the M 6.9 Miyagi earthquake.

There is a good agreement between the estimated (solid) and observed (dashed) source and ground-motion parameters for the three analyzed events (Fig. 6). In all six cases, reasonably well predicted parameters are available between 2.5 and around 10 s before PGV is observed. The strongest shaking in

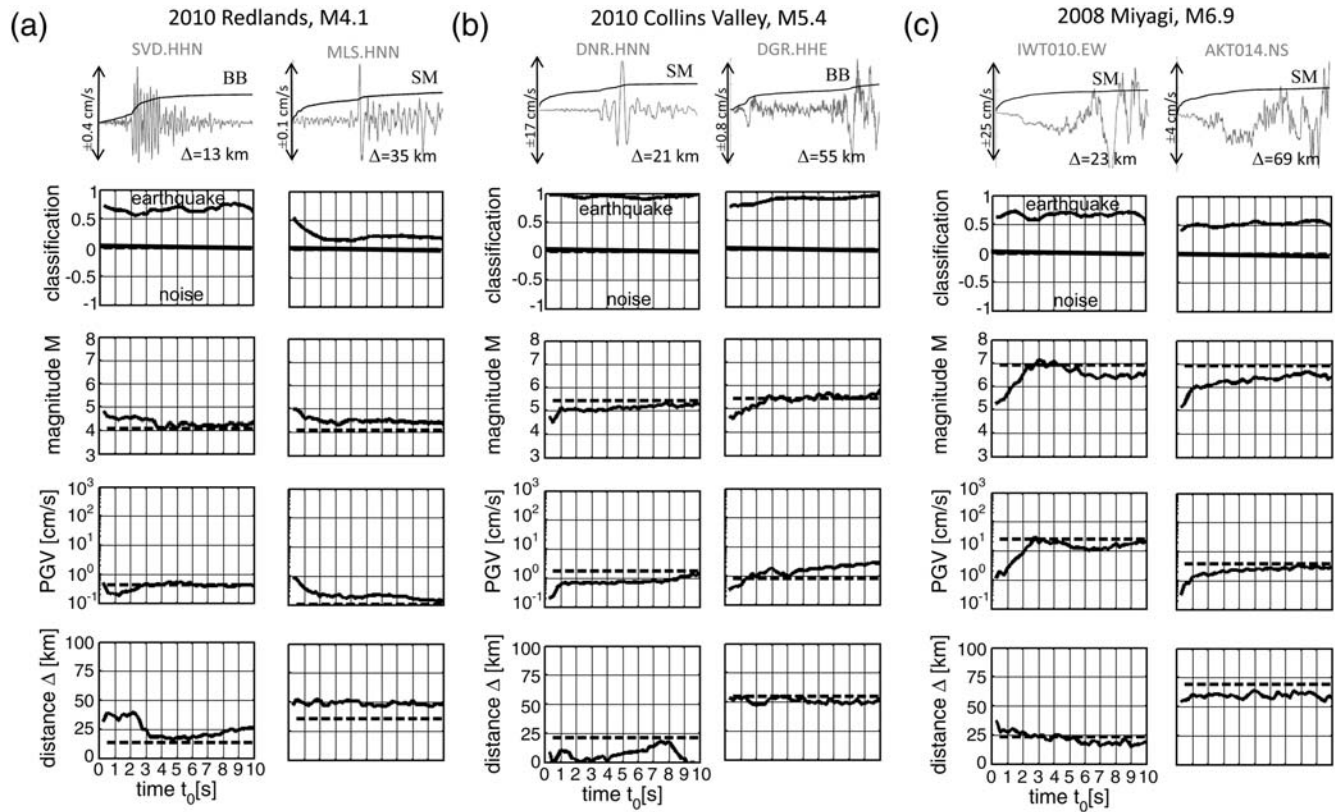


Figure 6. Demonstration of PreSEIS On-site for three earthquakes from the test dataset: (a) the 2010 M 4.1 Redlands and (b) the 2010 M 5.4 Collins Valley earthquakes in southern California, and (c) the 2008 M 6.9 Miyagi earthquake in Japan. For each earthquake, we show the results at two stations at distinct epicentral distances Δ . From top to bottom the panels show the corresponding seismic records and results for classification (output should be +1 because Δ [km] $\leq 15M$), magnitude M , peak ground velocity (PGV), and distance Δ . There is usually a good agreement between the estimates (solid lines) and the observed parameters (dashed lines) between 2.5 and around 10 s before PGV is observed. The initial magnitude estimates at $t_0 = 0.25$ s are almost the same for all three earthquakes, indicating that this time window is insufficient to resolve the size of the ongoing earthquakes.

these examples was observed at station IWT010. PreSEIS On-site recognized around 2 s after P -wave detection that PGV at this site would exceed 10 cm/s, offering a warning time of around 6 s prior to very strong shaking. Parameters estimated from data at stations at larger epicentral distances Δ usually require more time for convergence, in particular if the events have large magnitudes. We also expect longer warning times for these events, so that a warning could still be issued before strong shaking at most user sites occurs.

Discussion

PreSEIS On-site uses three-component acceleration, velocity, and displacement waveforms of a single BB or SM sensor to rapidly estimate earthquake source and ground-motion parameters for EEW. The trade-off between magnitudes and distances can be solved this way, because $\ddot{u}(t)$, $\dot{u}(t)$, and $u(t)$ show distinct dependencies on M and Δ (see also Yamada *et al.*, 2009). The three time series characterize different frequency bands of seismic ground motions: $\ddot{u}(t)$ is most sensitive to high frequencies (> 3 Hz), $\dot{u}(t)$ is most sensitive to midfrequencies (1 to 3 Hz), and $u(t)$ is most sensitive to low frequencies (< 1 Hz), and

$\dot{u}(t)$ is most sensitive to the midfrequency range (1 to 3 Hz). We expect that as an alternative to integrating and differentiating the recorded time series, we could use filters and consider different frequency bands, but such effort is beyond the scope of this study. Using integrated and differentiated time series has the advantage that the frequency spectra are simply weighted with ω or ω^2 , respectively, that is, our approach is scale-free, while (other) filtering approaches introduce one or more cutoff frequencies.

The majority of algorithms that have been proposed for EEW need to be calibrated to the seismic data of a particular area of interest, taking into account the regional differences in the seismic wave propagation and sometimes the earthquake source (e.g., Wu and Kanamori, 2005; Wu *et al.*, 2007; Cua *et al.*, 2009; Brown *et al.*, 2011). This procedure is problematic, because an adequate database containing the records of both moderate and large earthquakes over a wide range of source-to-site distances is unavailable for most seismic-active regions around the world, leading to large data gaps and thus uncertainties in the algorithms. Some authors have suggested filling these gaps with simulated waveforms (e.g., Böse *et al.*, 2008; Zollo *et al.*, 2009; Oth, Böse, *et al.*,

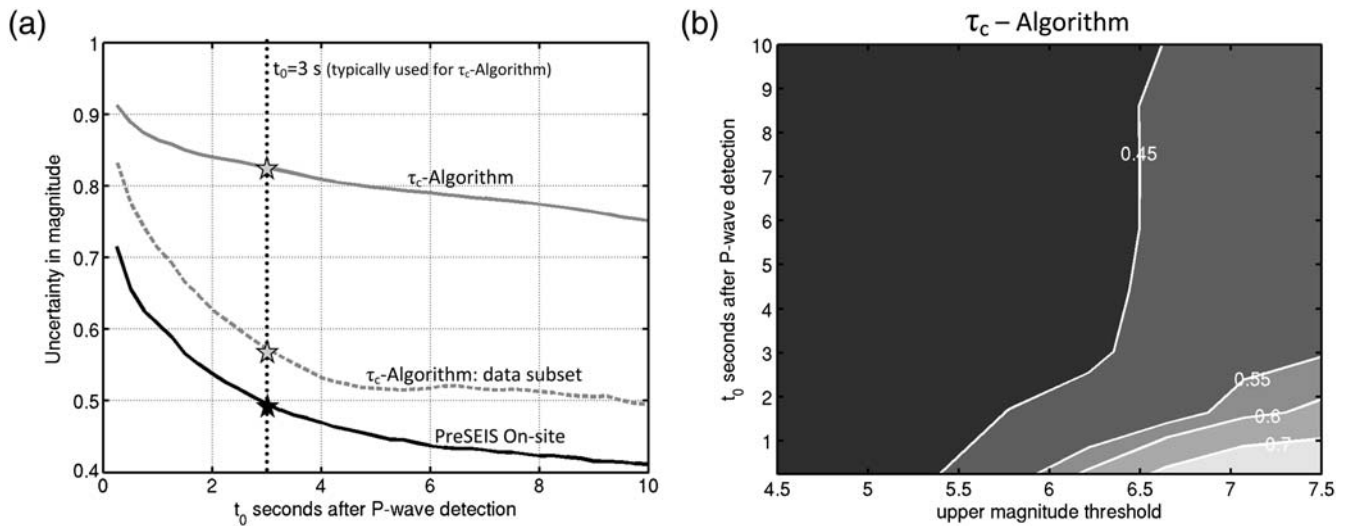


Figure 7. (a) Comparison of the results obtained from PreSEIS On-site and the τ_c algorithm. (b) Similar trends in the error distributions for both algorithms indicate some general nature of the predictability of earthquake ruptures and magnitudes using a limited time window of seismic data (compare with Figs. 3 and 5). The errors in the τ_c algorithm are around 60% higher than in PreSEIS On-site.

2010) or developing EEW systems based on theoretical considerations (Böse and Heaton, 2010), but it is questionable how reliable these algorithms will perform during a real major earthquake. Our approach of using global datasets of broadband (BB) and strong-motion (SM) records of small to major earthquakes and searching for the common characteristics of these events may be more reliable.

The magnitude-frequency distribution of events in the training dataset has impact on the absolute prediction errors of PreSEIS On-site, but the relative magnitude-time dependency of these errors remains almost unaffected (Fig. 4). This is an important observation implying that ANNs learn to some extent the *a priori* probabilities of the occurrence of earthquakes of different magnitudes from the training data. Alternatively, we may use a uniformly distributed dataset for the training and apply the *a priori* probabilities (such as the Gutenberg–Richter power-law statistics) afterward. By doing so, we can realize a fully Bayesian approach, similar to the virtual seismologist (VS) algorithm for EEW (Cua and Heaton, 2007). This method considers both the likelihood function (that describes how likely it is to observe a certain output, e.g., M , for a given input, e.g., IAA_i , IAV_i , and IAD_i), and the *a priori* probabilities (that are a measure for the frequency of the occurrence of a certain output, e.g., M) to determine probabilistic estimates.

SM sensors have two major advantages over BB sensors for EEW: (1) wave amplitudes stay on scale, that is, they are not clipped during strongest seismic shaking typically observed during large earthquakes ($M > 6.5$), which are most damaging; (2) SM sensors are usually less costly than BB instruments. On the other hand, the dynamic range of SM sensors is more limited. Strong-motion instruments are not as sensitive to ground motions as BB sensors that provide data for small and moderate earthquakes at moderate dis-

tances. These events usually do not cause structural damage (aside from the possible failure of weakened structures during aftershocks; Bakun *et al.* (1994)), but they occur more frequently than large earthquakes and can be used to optimize real-time EEW algorithms.

Furthermore, SM sensors are accelerometers; that is, to obtain information on displacement, which is important to calculate the static offset during an earthquake, the recorded acceleration time series needs to be double-integrated. Double integration, however, often results in significant long-period errors caused by very small linear baseline trends (e.g., caused by a small tilt of the sensor or thermal drift), which are difficult to isolate and remove from current strong-motion records (e.g., Clinton and Heaton, 2002). This can seriously distort the resultant displacement series (e.g., Boore, 2001).

Our results suggest that the magnitudes of large earthquakes require more time to be determined than magnitudes of small events (Fig. 4). Because ANNs have a great degree of flexibility and usually provide models with a very good or maybe even the best possible reproduction of the desired input-output mapping (at least for the training dataset), we suspect that these findings are not limited to PreSEIS On-site, but provide important insight into the predictability of earthquake ruptures in general.

The uncertainty in the magnitude prediction of large earthquakes decays slower with time than for small earthquakes because of their longer rupture durations. The systematic trends in the errors (Fig. 4) possibly reflect the ambiguity of the seismic data, which do not allow for a clear discrimination of small and large events at the very beginning of the rupture process. Kanamori (2005) suggests to set $t_0 = 3$ s, which corresponds to the typical rupture duration of an M 6 to M 6.5 earthquake, that is, within 3 s we should

be capable of deciding whether a detected earthquake is smaller or larger than M 6.5.

Comparison with the τ_c Algorithm

The input parameters of PreSEIS On-site in equations (1) to (3) show striking similarities with the period parameter τ_c , which was proposed by Kanamori (2005) for on-site early warning as an extension of a method developed by Nakamura (1988) and Allen and Kanamori (2003). The τ_c algorithm allows for a quick estimation of magnitude M of an earthquake based on the (assumed) log-linear relationship

$$M = a \log(\tau_c) + b, \quad (5)$$

where the coefficients a and b are specific for the earthquakes in a given region (e.g., Wu *et al.*, 2007). The τ_c parameter is defined by $\tau_c = \frac{2\pi}{r}$ with

$$r = \frac{\int_0^{t_0} \dot{u}_{UD}^2(t) dt}{\int_0^{t_0} u_{UD}^2(t) dt}, \quad (6)$$

where t_0 is usually set to 3 s. Using Parseval's theorem, Kanamori (2005) showed that $r \approx \frac{4\pi^2 \int_0^\infty f^2 |\hat{u}(f)|^2 df}{\int_0^\infty |\hat{u}(f)|^2 df} = 4\pi^2 \langle f^2 \rangle$, where $\hat{u}(f)$ is the frequency spectrum of $u(t)$, and $\langle f^2 \rangle$ is the average of f^2 weighted by $|\hat{u}(f)|^2$. The τ_c parameter is thus a measure of the effective period of ground shaking (Kanamori, 2005).

The comparison of equations (1) to (3) with equation (6) reveals similarities between τ_c and the PreSEIS On-site input. The main differences are that (1) the τ_c algorithm uses the squared amplitudes of $u(t)$ and $\dot{u}(t)$, while PreSEIS On-site uses absolute values (PreSEIS On-site using squared amplitudes was also tested, but no significant changes in the algorithm performance were detected); (2) PreSEIS On-site uses $\ddot{u}(t)$ in addition to $\dot{u}(t)$ and $u(t)$; (3) PreSEIS On-site uses three-component data; (4) PreSEIS On-site considers seismic site effects; (5) PreSEIS On-site does not assume a log-linear relationship of the seismic observations and M (equation 5), but is completely data-driven; and (6) PreSEIS On-site does not restrict the time window of integration to $t_0 = 3$ s, but starts at $t_0 = 0.25$ s after the earthquake detection and then continuously updates its predictions every 0.25 s.

To compare the performance of the two algorithms, we determine the τ_c – M relationship for the same dataset as used for the training and testing of PreSEIS On-site with $t_0 = 0.25$ to 10.0 s. For each earthquake we determine the median τ_c value taken over the values at all recording stations (BB and SM sensors) and determine at each time window t_0 the coefficients a and b in equation (5) from least-squares regression. The standard deviations σ of the obtained error distributions are shown in Figure 7a. Again the uncertainties in the predicted magnitudes decay with increasing t_0 , but are $\sim 60\%$ higher than for PreSEIS On-site (see

also Fig. 3). We apply a filter criterion (Böse *et al.*, 2009) to identify and remove records whose data quality might be insufficient for the τ_c algorithm. After that, the errors reduce significantly (Fig. 7a, gray dashed line), but are still more than 15% higher than for PreSEIS On-site.

Similar to Figure 5 we determine the standard deviations σ of the residuals for all events up to a certain upper magnitude threshold (Fig. 7b). Again we find that the magnitudes of large earthquakes require more time to stabilize than magnitudes of small events. The uncertainties in Figures 5 and 7a cannot be directly compared with each other, because Figure 7b refers to the event median. Taking the event median is required to stabilize the predictions by the τ_c algorithm (see, e.g., Allen and Kanamori, 2003). Typically, the parameters a and b in equation (5) are optimized for the earthquakes in a specific region and over a smaller distance range (e.g., Wu and Kanamori, 2005; Wu *et al.*, 2007; Böse, Hauksson, Solanki, Kanamori, and Heaton, 2009) than was done in this study. The observed magnitude errors of the τ_c algorithm are thus usually smaller than shown in Figure 7.

Conclusions and Outlook

We developed and tested a new algorithm for earthquake early warning (EEW) that uses three-component broadband (BB) or strong-motion (SM) waveforms recorded at a single sensor. Based on artificial neural networks (ANNs), PreSEIS On-site classifies earthquake/noise and near/far source events, as well as estimates the moment magnitude M , epicentral distance Δ , and the peak ground velocity (PGV) at the site of observation. First estimates become available at $t_0 = 0.25$ s after the P pick and are regularly updated. The magnitude-distance trade-off is solved from usage of information derived from the acceleration, velocity, and displacement waveform time series, $\ddot{u}(t)$, $\dot{u}(t)$, and $u(t)$, that show distinct dependencies on M and Δ . We find that $u(t)$ and $\dot{u}(t)$ are most relevant for the estimation of M and PGV, while $\ddot{u}(t)$ is important for the earthquake/noise discrimination.

PreSEIS On-site overcomes the limitations of the former (network-based) PreSEIS algorithm (Böse *et al.*, 2008) by being faster and network-independent. The algorithm has been tested successfully with a large dataset of recorded waveforms from different tectonic settings (California, Japan, and Taiwan). PreSEIS On-site does not replace the original PreSEIS algorithm that clearly has important features arising from the usage of multiple sensors. The implementation of PreSEIS On-site, however, is more straightforward, and the code is directly applicable (Appendix A).

At each time t_0 after P -wave detection, PreSEIS On-site uses the current (logarithmic) values of the integrated absolute amplitudes of the acceleration, velocity, and displacement waveforms, IAA_i , IAV_i and IAD_i ($i = \{EW, NS, UD\}$). Taking the logarithm of these values has a smoothing effect for larger amplitudes, including the seismic S wave (see, e.g., Fig. 6, top graphs). The temporal evolution of IAA_i , IAV_i , and IAD_i could give additional information

and help improving predictions. This shall be explored in future studies.

This study was based on a joint dataset of BB and SM records of 161 crustal earthquakes from California, Japan, and Taiwan with $3.1 \leq M \leq 7.6$ at epicentral distances of up to 115 km. The initial uncertainties of $M \pm 0.7$ and $\log(\text{PGV}) \pm 0.55$ for this dataset decrease with progressing time, revealing a trade-off between the reliability of warnings and remaining warning time. We found systematic trends in the prediction errors, such that the parameters for small earthquakes tend to be overestimated, while large events tend to be underestimated. These trends are more profound at the beginning of the rupture process and seem to be caused by the ambiguity of the waveform data from small and large earthquake at rupture initiation. We compared our approach with the τ_c algorithm (Kanamori, 2005) and found that the prediction errors of PreSEIS On-site are around 60% smaller. However, in general there is a limitation to the prediction accuracy an EEW system can provide if based on single-sensor observations.

We tested PreSEIS On-site also for subduction-zone events in Japan in the same magnitude range $3.1 \leq M \leq 7.6$ and found that our approach is not limited to crustal earthquakes (results are not shown here). However, very large subduction-zone earthquakes that are often accompanied by devastating tsunamis, such as the 2011 M 9.0 Tohoku earthquake in Japan, the 2010 M 8.8 Maule (Chile) earthquake, or the 2004 M 9.2 Sumatra–Andaman earthquake, pose a huge challenge to EEW. Due to long rupture durations of one or more minutes, these events clearly require updating procedures for magnitude estimations longer than 10 s as we used for crustal earthquakes in this study. Earthquakes with M 8.0 and larger might need some special treatment for EEW and will likely require the inclusion of other types of data, such as from a real-time Global Positioning System (e.g., Crowell *et al.*, 2009; Böse and Heaton, 2010; Hammond *et al.*, 2011).

Data and Resources

Broadband and strong-motion records used in this study were downloaded from (1) CISN, operated by the California Institute of Technology (Caltech), USGS Pasadena/Menlo Park, California Geological Survey, and UC Berkeley; (2) K-NET, operated by the Japanese National Research Institute for Earth Science and Disaster Prevention (NIED); and (3) TSMIP, operated by the Chinese Weather Bureau. The data for the 1999 Chi-Chi earthquake and its aftershocks were downloaded from the COSMOS Virtual Datacenter (www.cosmos-eq.org/, last accessed October 2011). The moment magnitudes M from the Global Centroid Moment Tensor Catalog (www.globalcmt.org/CMTsearch.html, last accessed October 2011) were used to create a consistent dataset rather than using the JMA magnitude M_{JMA} . The K-NET web site provided borehole data and shear-wave velocities used for soil classification (www.k-net.bosai.go.jp/, last accessed Octo-

ber 2011). The National Center for Research on Earthquake Engineering and the Chinese Weather Bureau (<http://geo.ncree.org.tw>, last accessed October 2011) provided the V_{S30} values. PreSEIS On-site can be obtained upon request.

Acknowledgments

This work is funded through contract G09AC00258 from USGS/ANSS to the California Institute of Technology (Caltech). This is contribution #10058 of the Seismological Laboratory, Geological and Planetary Sciences at Caltech. We would like to thank William H. Bakun and an anonymous reviewer for their helpful comments.

References

- Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces, *Bull. Seismol. Soc. Am.* **68**, 1521–1532.
- Allen, R. M., and H. Kanamori (2003). The potential for earthquake early warning in Southern California, *Science* **300**, 786–789.
- Allen, R. M., P. Gasparini, O. Kamigaichi, and M. Böse (2009). The status of earthquake early warning around the world: An introductory overview, *Seismol. Res. Lett.* **80**, no. 5, 682–693, doi [10.1785/gssrl.80.5.682](https://doi.org/10.1785/gssrl.80.5.682).
- Bakun, W. H., F. G. Fischer, E. G. Jensen, and J. VanSchaak (1994). Early warning system for aftershocks, *Bull. Seismol. Soc. Am.* **84**, no. 2, 359–365.
- Bishop, C. (1995). *Neural Networks for Pattern Recognition* Oxford, Clarendon Press, 482 pp.
- Böse, M. (2006). Earthquake early warning for Istanbul using artificial neural networks, *Ph.D. Thesis*, 181 pp., Karlsruhe University, Germany, <http://www.ubka.uni-karlsruhe.de/cgi-bin/psview?document=2006/physik/6>, last accessed October 2011.
- Böse, M., and T. H. Heaton (2010). Probabilistic prediction of rupture length, slip and seismic ground motions for an ongoing rupture: Implications for early warning for large earthquakes, *Geophys. J. Int.* **183**, no. 2, 1014–1030, doi [10.1111/j.1365-246X.2010.04774.x](https://doi.org/10.1111/j.1365-246X.2010.04774.x).
- Böse, M., E. Hauksson, K. Solanki, H. Kanamori, Y.-M. Wu, and T. H. Heaton (2009). A new trigger criterion for improved real-time performance of on-site earthquake early warning in southern California, *Bull. Seismol. Soc. Am.* **99**, no. 2-A 897–905, doi [10.1785/0120080034](https://doi.org/10.1785/0120080034).
- Böse, M., E. Hauksson, K. Solanki, H. Kanamori, and T. H. Heaton (2009). Real-time testing of the on-site warning algorithm in southern California and its performance during the July 29 2008 M_w 5.4 Chino Hills earthquake, *Geophys. Res. Lett.* **36**, L00B03, doi [10.1029/2008GL036366](https://doi.org/10.1029/2008GL036366).
- Böse, M., C. Ionescu, and F. Wenzel (2007). Earthquake early warning for Bucharest, Romania: Novel and revised scaling relations, *Geophys. Res. Lett.* **34**, L07302, doi [10.1029/2007GL029396](https://doi.org/10.1029/2007GL029396).
- Böse, M., F. Wenzel, and M. Erdik (2008). PreSEIS: A neural network based approach to earthquake early warning for finite faults, *Bull. Seismol. Soc. Am.* **98**, no. 1, 366–382, doi [10.1785/0120070002](https://doi.org/10.1785/0120070002).
- Boore, D. M. (2001). Effect of baseline corrections on displacements and response spectra for several recordings of the 1999 Chi-Chi, Taiwan, earthquake, *Bull. Seismol. Soc. Am.* **91**, no. 1, 199–211.
- Brown, H. M., R. M. Allen, M. Hellweg, O. Khainovski, D. Neuhauser, and A. Souf (2011). Development of the ElarmS methodology for earthquake early warning: Realtime application in California and offline testing in Japan, *Soil Dynam. Earthquake Eng.* **31**, 188–200, doi [10.1016/j.soildyn.2010.03.008](https://doi.org/10.1016/j.soildyn.2010.03.008).
- Clinton, J. F., and T. H. Heaton (2002). Potential advantages of a strong-motion velocity meter over a strong-motion accelerometer, *Seismol. Res. Lett.* **73**, no. 3, 332–342.
- Crowell, B. W., Y. Bock, and M. B. Squibb (2009). Demonstration of earthquake early warning using total displacement waveforms from real-time GPS networks, *Seismol. Res. Lett.* **80**, no. 5, 772–782, doi [10.1785/gssrl.80.5.772](https://doi.org/10.1785/gssrl.80.5.772).

- Cua, G., and T. Heaton (2007). The Virtual Seismologist (VS) method: A Bayesian approach to earthquake early warning, in *Earthquake Early Warning Systems*, P. Gasparini, G. Manfredi, and J. Zschau (Editors), Springer, New York, 85–132.
- Cua, G., M. Fischer, T. Heaton, and S. Wiemer (2009). Real-time performance of the Virtual Seismologist earthquake early warning algorithm in Southern California, *Seismol. Res. Lett.* **80**, no. 5, 740–747, doi [10.1785/gssrl.80.5.740](https://doi.org/10.1785/gssrl.80.5.740).
- Gutenberg, R., and C. F. Richter (1944). Frequency of earthquakes in California, *Bull. Seismol. Soc. Am.* **34**, 185–188.
- Hammond, W. C., B. A. Brooks, R. Bürgmann, T. Heaton, M. Jackson, A. R. Lowry, and S. Anandakrishnan (2011). Scientific value of real-time Global Positioning System data, *Eos Trans. AGU* **92**, no. 15, 125–132.
- Hilbring, D., T. Titzschkau, A. Buchmann, G. Bonn, F. Wenzel, and E. Hohnacker (2010). Earthquake early warning for transport lines, *Nat. Hazards*, doi [10.1007/s11069-010-9609-3](https://doi.org/10.1007/s11069-010-9609-3).
- Horiuchi, S., H. Negishi, K. Abe, A. Kamimura, and Y. Fujinawa (2005). An automatic processing system for broadcasting earthquake alarms, *Bull. Seismol. Soc. Am.* **95**, no. 2, 708–718.
- Kanamori, H. (2005). Real-time seismology and earthquake damage mitigation, *Annu. Rev. Earth Planet. Sci.* **33**, 195–214, doi [10.1146/annurev.earth.33.092203.122626](https://doi.org/10.1146/annurev.earth.33.092203.122626).
- Köhler, N. (2010). Real-time information from seismic networks, *Ph.D. Thesis*, 150 pp., Karlsruhe Institute of Technology, Germany, <http://digbib.ubka.uni-karlsruhe.de/volltexte/1000015555>, last accessed October 2011.
- Köhler, N., G. Cua, F. Wenzel, and M. Böse (2009). Rapid source parameter estimations of Southern California earthquakes using PreSEIS, *Seismol. Res. Lett.* **80**, no. 5, 748–754, doi [10.1785/gssrl.80.5.748](https://doi.org/10.1785/gssrl.80.5.748).
- Leach, R., and F. Dowla (1996). Earthquake early warning system using real-time signal processing, in *IEEE Workshop on Neural Networks for Signal Processing*, Keihanna, Kyoto, Japan, 4–6 September 1996.
- Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares, *Q. J. Appl. Math.* **2**, 164–168.
- Nakamura, Y. (1988). On the urgent earthquake detection and alarm system (UrEDAS), *Proc. of the 9th World Conference on Earthquake Engineering*, Tokyo–Kyoto, Japan.
- National Earthquake Hazards Reduction Program (1994). Recommended provisions for seismic regulations for new buildings, *Federal Emergency Management Agency Rept. FEMA 222A*, Washington, D.C., 290 pp.
- Olson, E. L., and R. M. Allen (2005). The deterministic nature of earthquake rupture, *Nature* **438**, 212–215.
- Oth, A., D. Bindi, S. Parolai, and D. Di Giacomo (2010). Earthquake scaling characteristics and the scale-(in)dependence of seismic energy-to-moment ratio: Insights from KiK-net data in Japan, *Geophys. Res. Lett.* **37**, L19304, doi [10.1029/2010GL044572](https://doi.org/10.1029/2010GL044572).
- Oth, A., M. Böse, F. Wenzel, N. Köhler, and M. Erdik (2010). Evaluation and optimization of seismic networks and algorithms for earthquake early warning—The case of Istanbul (Turkey), *J. Geophys. Res.* **115**, B10311, doi [10.1029/2010JB007447](https://doi.org/10.1029/2010JB007447).
- Rumelhart, D., G. Hinton, and R. Williams (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*, vol. **1**, MIT Press, Cambridge, 564 pp.
- Rydelek, P., and S. Horiuchi (2006). Earth science: Is the earthquake rupture deterministic? *Nature* **442**, doi [10.1038/nature04963](https://doi.org/10.1038/nature04963).
- Rydelek, P., C. Wu, and S. Horiuchi (2007). Comment on “Earthquake magnitude estimation from peak amplitudes of very early seismic signals on strong motion records” by Aldo Zollo, Maria Lancieri, and Stefan Nielsen, *Geophys. Res. Lett.* **34**, L20302, doi [10.1029/2007GL029387](https://doi.org/10.1029/2007GL029387).
- Wald, D. J., V. Quitoriano, T. H. Heaton, H. Kanamori, C. W. Scrivner, and C. B. Worden (1999). TriNet ShakeMaps: rapid generation of instrumental ground motion and intensity maps for earthquakes in Southern California, *Earthquake Spectra* **15**, 537–556.
- Wills, C. J., M. D. Petersen, W. A. Bryant, M. S. Reichle, G. J. Saucedo, S. S. Tan, G. C. Taylor, and J. A. Treiman (2000). A site conditions map for California based on geology and shear wave velocity, *Bull. Seismol. Soc. Am.* **90**, no. 6b, 187–208.
- Wu, Y.-M., and H. Kanamori (2005). Experiment on an on-site early warning method for the Taiwan early warning system, *Bull. Seismol. Soc. Am.* **95**, 347–353.
- Wu, Y.-M., and T. L. Teng (2002). A virtual sub-network approach to earthquake early warning, *Bull. Seismol. Soc. Am.* **92**, 2008–2018.
- Wu, Y.-M., H. Kanamori, R. M. Allen, and E. Hauksson (2007). Determination of earthquake early warning parameters, τ_c and P_d , for southern California, *Geophys. J. Int.* **170**, 711–717, doi [10.1111/j.1365-246X.2007.03430.x](https://doi.org/10.1111/j.1365-246X.2007.03430.x).
- Yamada, T., and S. Ide (2008). Limitation of the predominant-period estimator for earthquake early warning and the initial rupture of earthquakes, *Bull. Seismol. Soc. Am.* **98**, no. 6, 2739–2745, doi [10.1785/0120080144](https://doi.org/10.1785/0120080144).
- Yamada, M., T. Heaton, and J. Beck (2007). Real-time estimation of fault rupture extent using near-source versus far-source classification, *Bull. Seismol. Soc. Am.* **97**, no. 6, 1890–1910, doi [10.1785/0120060243](https://doi.org/10.1785/0120060243).
- Yamada, M., A. H. Olsen, and T. H. Heaton (2009). Statistical features of short-period and long-period near-source ground motions, *Bull. Seismol. Soc. Am.* **99**, no. 6, 3264–3274, doi [10.1785/0120090067](https://doi.org/10.1785/0120090067).
- Zollo, A., G. Iannaccone, M. Lancieri, L. Cantore, V. Convertito, A. Emolo, G. Festa, F. Gallovič, M. Vassallo, C. Martino, C. Satriano, and P. Gasparini (2009). Earthquake early warning system in southern Italy: Methodologies and performance evaluation, *Geophys. Res. Lett.* **36**, L00B07, doi [10.1029/2008GL036689](https://doi.org/10.1029/2008GL036689).
- Zollo, A., M. Lancieri, and S. Nielsen (2006). Earthquake magnitude estimation from peak amplitudes of very early seismic signals on strong motion records, *Geophys. Res. Lett.* **33**, L23312, doi [10.1029/2006GL027795](https://doi.org/10.1029/2006GL027795).

Appendix A

Two-Layer-Feed-Forward Neural Networks

Two-layer-feed-forward (TLFF) neural networks are composed of simple processing units (called neurons), which are arranged in input layers, hidden layers, and output layers, respectively. The neurons of different layers are connected to each other and exchange information (Fig. 1b). The importance of the links between different neurons is controlled by weight parameters that are determined during the training of the ANNs using a set of example patterns (e.g., Bishop, 1995).

The output y of a TLFF neural network is calculated from the input vector \mathbf{x} and the weights in the input and hidden layers, $w_{ji}^{(1)}$ and $w_j^{(2)}$:

$$y = \sum_{j=0}^J w_j^{(2)} g\left(\sum_{i=0}^I w_{ji}^{(1)} x_i\right), \quad (\text{A1})$$

with $x_0 = 1$. To encompass nonlinear behavior of the network, we apply the logistic activation function $g(\arg) = 1/[1 + \exp(-\arg)]$.

For each output parameter y (M , Δ , and PGV and classification), we build one TLFF network with $I = 10$ input, $J = 15$ hidden, and $K = 1$ output units, which gives us a total of 181 network weights. The training occurs from a set of example patterns with known input and output values. We adopt the Levenberg optimization method (Levenberg,

1944) combined with a back-propagation algorithm (e.g., Rumelhart *et al.*, 1986) for iteratively updating the network weights to decrease the mean absolute errors between the observed and desired outputs: $M_{\text{err}} = M_{\text{obs}} - M_{\text{pred}}$, $\log(\Delta_{\text{err}}) = \log(\Delta_{\text{obs}}) - \log(\Delta_{\text{pred}})$, and $\log(\text{PGV}_{\text{err}}) = \log(\text{PGV}_{\text{obs}}) - \log(\text{PGV}_{\text{pred}})$. To avoid the overfitting of the ANNs to the training dataset (at the expense of the desired generalization capability for new data), we use an independent validation dataset (which is used neither for training nor for testing) and stop the iterative weight update, once the errors for these data start increasing. The training of each TLFF network usually requires less than 50 iterations, which takes roughly 1 s on a common PC (see Böse, 2006 and Böse *et al.*, 2008 for more details).

Appendix B

Dataset

The CISN dataset comprises 703 three-component BB and 677 three-component SM records of 107 crustal earthquakes with $3.1 \leq M \leq 7.1$ at epicentral distances of 1.2 to 105 km that occurred in northern and southern California between 1990 and 2010. The set includes, for example, the 1994 M 6.7 Northridge and the 1999 M 7.1 Hector Mine earthquakes. We employ the V_{S30} values at the CISN stations taken from Wills *et al.* (2000) that are also used in the Californian ShakeMaps (Wald *et al.*, 1999).

The K-NET dataset comprises 938 three-component SM records of 48 earthquakes with $4.8 \leq M \leq 7.3$ at epicentral

distances of 3.7 to 110 km, including, for example, the 2003 M 7.3 Tokachi–Oki and the 2008 M 6.9 Iwate–Miyagi earthquakes. To obtain a consistent dataset, we replaced the JMA magnitude M_{JMA} with the moment magnitudes M from the Global Centroid Moment Tensor Catalog (see Data and Resources) if available; for smaller events, we assumed that $M \approx M_{\text{JMA}}$ (Oth, Bindi, *et al.*, 2010). For the soil classification at the K-NET station sites we used close-by borehole data and averaged the shear-wave velocities down to their full depth (see Data and Resources). Because most boreholes have depths of 20 m only, we assumed that $V_{S30} \approx V_{S20}$.

The Taiwanese dataset includes 113 three-component SM records of 6 earthquakes with $5.0 \leq M \leq 7.6$ at epicentral distances of 5.2 km to 115 km. The set comprises records of the 1999 M 7.6 Chi-Chi earthquake and its five strongest aftershocks. The data were downloaded from the COSMOS database (see Data and Resources). V_{S30} values were obtained from the National Center for Research on Earthquake Engineering and the Chinese Weather Bureau (see Data and Resources).

Seismological Laboratory
California Institute of Technology
1200 E. California Blvd.
Mail Code 252–21
Pasadena, California 91125
mboese@caltech.edu
hauksson@caltech.edu
heaton@caltech.edu

Manuscript received 18 May 2011